DOCUMENT RESUME

ED 408 322 TM 026 575

AUTHOR Bode, Rita K.
TITLE Measuring Feedback-Seeking Modes: An Alternative to Composite Scores.
PUB DATE Mar 97
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Faculty; *Feedback; Higher Education; Measurement Techniques; *Rating Scales; Scaling; *Surveys; *Test Construction
IDENTIFIERS BIGSTEPS Computer Program; Calibration; *Composite Scores; Monitoring; *Rasch Model

ABSTRACT

An alternative to the use of traditional composite scales in creating scales from survey items was developed, using feedback seeking as an example. It is proposed that much more information can be obtained through Rasch techniques about feedback-seeking behaviors related to teaching. Data are from the New Faculty Project involving tenure-track faculty hired in 1991 and 1992 at five institutions in a consortium. Responses from 165 faculty members indicated 9 activities through which faculty members sought feedback about job performance. Instead of creating composite scores in the traditional way, researchers Rasch-calibrated responses using the BIGSTEPS, version 2.65 program. Results show that new faculty members are more likely to monitor the behavior of others than to make inquiries about job performance. Faculty members could be distinguished by their feedback-seeking modes. More information on the relationship between monitoring and feedback and the frequency of feedback was obtained than could be obtained through the traditional use of composite scores. Composite scores made it possible to distinguish those who used a feedback-seeking modes more or less frequently, but not to investigate patterns of feedback seeking. (Contains two tables, four figures, and four references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 408 322

Measuring Feedback-Seeking Modes:

An Alternative to Composite Scores

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
RITA BODE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Rita K. Bode

Rehabilitation Institute of Chicago

March 1997

Paper presented at the Annual Meeting of the
American Educational Research Association, Chicago IL

TM026575

BEST COPY AVAILABLE

Measuring Feedback-Seeking Modes: An Alternative to Composite Scores

This study illustrates an alternative to the use of traditional composite scores in creating scales from survey items. Traditional composite scores are known to be limited in the amount of information they provide about the individuals measured and the use of Rasch analysis is proposed to more fully describe the behaviors or activities being measured (Wright & Masters, 1982).

The variable dealt with in this study is mode of feedback-seeking. Newcomers in many employment situations deal with anxiety by seeking information on what is expected of them and how well they are doing in their new jobs. Rather than just passively receiving information, these employees seek out information in general and feedback on their performance in specific. Three modes in which information can be sought have been described by Morrison (1993) as follows: consulting written documents for information about expectations, monitoring the behavior of others to obtain information, and making inquiries of colleagues and supervisors to obtain feedback related to performance.

In previous research, Morrison (1993) created a scale consisting of seven items and used it to compare the feedback-seeking modes of newly-hired staff accountants. The response scale represented the frequency with which each behavior was used: from 1 (never) to 7 (several times a day). The items consisted of the following: 1) ask your direct supervisor, 2) ask a more experienced staff accountant, 3) ask another new staff accountant, 4) pay attention to how others behave, 5) socialize with people in the firm in order to learn how they behave and what they value, 6) observe what behaviors are rewarded and use this as a clue to what is desirable or expected, and 7) consult memos, annual reports, or other written material. She used factor analysis to determine the number of scales to be created: the first three items comprised the inquiry scale, the second three comprised the monitor scale, and consult written documents was measured by the last item. A composite score was created for the two scales with Cronbach's alpha ranging from .60 to .65. Morrison compared the average scale/item scores for feedback about performance and found that new hires monitored significantly more frequently than made inquiries or consulted written documents. In this study, Rasch techniques are used to show how more information can be obtained on feedback-seeking behaviors related to teaching (logistics, expectations, performance, etc.).

METHODS

The data for this study came from the New Faculty Project survey that was sent to tenure-track faculty hired in 1991 and 1992. (The New Faculty Project was supported by the Office of Educational Research and Improvement (OERI) of the United States Department of Education through the National Center on Postsecondary Teaching, Learning, and Assessment (NCTLA), an effort of a consortium of universities located at Pennsylvania State University. The New Faculty Project, based at Northwestern University, was under the direction of Robert J. Menges, Professor of Education and Social Policy.) The measure of feedback seeking was developed for a study of feedback seeking and receiving in newly-hired faculty (Menges, Bode, Reyes, & Letwat, 1996).

Sample

The data for this study came from the New Faculty Project survey that was sent to faculty from two cohorts who were new to their institutions. These faculty came from five institutions: two liberal arts colleges, one community college district, one comprehensive university, and one research university. These faculty were surveyed for three years; responses used in this study came from year three when 252 surveys were mailed and 176 returned. While faculty of all ranks were included in the survey, the sample used in this study were unranked or below the rank of full professors. The data used in this study comes from 165 faculty who responded to at least one item in this set. Of these faculty, 79 were male and 86 were female; 15 were instructors, 80 were assistant professors, 20 were associate professors, 7 were full professors, and the remainder had other ranks or came from institutions with no ranks; 58 had minimal teaching experience (the equivalent of less than 2 years), 55 had moderate teaching experience (the equivalent of between 2 and 6 years), and 46 had extensive teaching experience (the equivalent of more than 7 years). The disciplines these faculty came from were varied but were more likely in the humanities or professions than in the natural or social sciences.

Instruments and Instrumentation

Survey items to assess feedback-seeking mode were adapted from Morrison (1993b). These items asked about the frequency with which new hires sought information from various sources, slightly modified from the earlier application in a business setting to one more appropriate for a postsecondary setting. A source of feedback was added to the item: *ask an administrator other than your chair*. Feedback-seeking mode thus was measured by an item that included nine activities by which faculty sought feedback. "Other" (option i) was not included in the scaling. The survey item on which this variable was based is as follows:

> *Think about your teaching and research/creative activities in the past academic year. In trying to figure out how well you were performing your work, how frequently in general did you do the things listed below. Please use the following scale:*
>
> *0 = never*
> *1 = once a month*
> *2 = few times a month*
> *3 = once a week*
> *4 = few times a week*
> *5 = once a day*
> *6 = few times a day*
>
> *a. ask your chair for feedback*
> *b. ask a more experienced senior colleague*
> *c. ask another junior faculty member*
> *d. ask an administrator other than your chair*
> *e. pay attention to how colleagues behave*
> *f. socialize with colleagues to learn how they behave and what they value*
> *g. observe what behaviors are rewarded and use this as a clue to what is desirable or expected*
> *h. consult memos, handbooks, or other written documents*
> *i. other*

Analysis

Instead of creating composite scores a la Morrison, the items were Rasch-calibrated using BIGSTEPS, version 2.65 (Linacre, 1992). Initially all items were calibrated to detect misfitting items; in subsequent analyses of fitting items only, step disorder was examined. The criterion for item misfit was set at a mean square infit statistic greater than 1.3 and step calibrations were examined to identify step disorder. Misfitting items were deleted from the set and step disorder was eliminated by collapsing adjacent categories in which the rating scale was used inconsistently. The standard item map provided in Table 12 of the BIGSTEPS output contains calibrations for a single value in the middle of the response scale (in this case, category 3, once a week). To more fully describe faculty who fell at various positions along the continuum, this item map was expanded using calibrations for each step in the response scale that are provided in Table 2.2 of the BIGSTEPS output.

RESULTS

The initial calibration indicated that one item, *consult written documents*, misfit the model. While faculty responses to the remaining items fit a general construct of feedback seeking, those for the misfitting item were unexpected--either unexpectedly frequent or unexpectedly infrequent--based on responses to the set of items. This item was thus eliminated from the measure. Step disorder was found at the top of scale; there were inconsistencies in the way that faculty used the categories *once a day* and *a few times a day* so these two responses were collapsed. Using a six-step response scale, the remaining six items created a useful measure of feedback-seeking with acceptable item infit and person separation reliability. The summary statistics for the final calibration are presented in Table 1. The calibrations and fit statistics for each item in measure order are presented in Table 2.

Table 1

SUMMARY OF 165 MEASURED (NON-EXTREME) PERSONS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	10.9	6.9	-1.07	.45	.99	-.3	1.03	-.2
S.D.	5.5	.3	1.03	.12	.69	1.1	.79	1.1

MODEL RMSE	.46	ADJ.SD	.92	SEPARATION	1.99	PERSON RELIABILITY	.80
REAL RMSE	.55	ADJ.SD	.87	SEPARATION	1.58	PERSON RELIABILITY	.71
S.E. OF PERSON MEAN	.08						
WITH 6 EXTREME PERSONS = 171 PERSONS							
MODEL RMSE	.53	ADJ.SD	1.08	SEPARATION	2.04	PERSON RELIABILITY	.81
REAL RMSE	.61	ADJ.SD	1.04	SEPARATION	1.72	PERSON RELIABILITY	.75

MINIMUM EXTREME SCORE: 6 PERSONS

Table 1, Continued

SUMMARY OF 7 MEASURED (NON-EXTREME) ITEMS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	257.1	162.6	.00	.09	1.00	-.1	1.02	.0
S.D.	126.9	1.8	1.08	.03	.19	1.6	.14	1.1

MODEL RMSE .10 ADJ.SD 1.08 SEPARATION 11.00 ITEM RELIABILITY .99
REAL RMSE .10 ADJ.SD 1.08 SEPARATION 10.40 ITEM RELIABILITY .99
S.E. OF ITEM MEAN .44

DELETED: 1 ITEMS

SUMMARY OF MEASURED STEPS

CATEGORY LABEL	STEP VALUE	OBSERVED COUNT	AVGE MEASURE	OUTFIT MNSQ	OB/EX FIT	STEP MEASURE	STEP ERROR	EXPECTED SCORE MEASURES STEP-.5	AT STEP	STEP+.5	THURSTONE THRESHOLD
0	0	328	-2.50	.87	1.06	NONE			(-3.08)	-2.28	
1	1	310	-1.32	.86	1.02	-1.81	.09	-2.28	-1.37	-.76	-2.03
2	2	218	-.50	1.01	1.02	-.54	.08	-.76	-.30	.08	-.67
3	3	125	.01	1.28	1.14	.33	.09	.08	.45	.84	.14
4	4	106	.53	1.38	1.02	.47	.11	.84	1.34	2.12	.73
5	5	51	1.25	1.70	.81	1.56	.17	2.12	(2.87)		1.83
						modal			mean		median

Table 2
ITEMS STATISTICS: MEASURE ORDER

ENTRY NUM	RAW SCORE	COUNT	MEASURE	ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTBIS	ITEMS
4	51	163	2.14	.16	1.18	1.1	1.20	.9	.33	TCHG--ASK ADMINISTRATOR
1	140	159	.73	.10	.85	-1.2	.99	-.1	.51	TCHG--ASK CHAIR
3	206	162	.16	.09	1.20	1.6	1.09	.7	.50	TCHG--ASK OTHER JUNIOR FACULTY
2	237	164	-.03	.08	.68	-3.1	.80	-1.8	.58	TCHG--ASK SENIOR COLLEAGUES
7	368	165	-.85	.08	1.18	1.6	1.21	1.8	.44	TCHG--OBSERVE WHAT IS REWARDED
6	377	163	-.94	.08	1.03	.3	.99	-.1	.53	TCHG--SOCIALIZE WITH COLLEAGUE
5	421	162	-1.21	.08	.90	-1.0	.89	-1.1	.57	TCHG--WATCH HOW COLLEAGUES BEH
MEAN	257.	163.	.00	.09	1.00	-.1	1.02	.0		
S.D.	127.	2.	1.08	.03	.19	1.6	.14	1.1		

The original item map is presented in Figure 1. This map shows that faculty in general are more likely to monitor the behavior of others than make inquiries of others which confirms the results obtained by Morrison (1993). However, additional information concerning feedback-seeking is also available from this analysis. For faculty who inquire, the frequency with which they monitor the behavior of others was essentially the same regardless of behavior being monitored but the frequency with which inquiries are made differed across source: faculty were least likely to ask administrators for feedback concerning their teaching performance, more likely to ask their chairs, and most likely to ask their colleagues (either junior or senior faculty).

While this map did provide more information about feedback-seeking than obtained in the previous research, results from Table 2.2 were used to expand the item map and more fully describe feedback-seeking. Figure 2 shows the calibrations for each step in the response scale for each item. Using these calibrations it was possible to illustrate how frequently feedback was sought by creating an expanded item map with descriptors for each response stacked by source.

The expanded item map was used to identify four types of feedback-seekers by referencing the item descriptors. The four types are: faculty who neither monitor nor inquire; faculty who monitor only; faculty who monitor frequently but inquire infrequently; and faculty who monitor and inquire frequently. Specifically, faculty who "neither monitor nor inquire" never seek feedback from any source; they may not feel the need for feedback or possibly were disappointed in the feedback they previously obtained. Faculty who "monitor only" monitor the behavior of others about once a month; they may feel the need for feedback but fear asking for it directly. Faculty who "monitor frequently but inquire infrequently" monitor the behavior of others a few times a month but only ask colleagues once a month and never ask chairs or administrators; they may feel the need for feedback but are only comfortable in seeking it from fellow faculty members. Faculty who "monitor and inquire frequently" monitor the behavior of others from once a week to at least once daily, ask colleagues for feedback from a few times a month to at least once daily, ask their chairs for feedback from once a month to a few times a day, and ask administrators for feedback never to once a week; they feel the need for feedback and appear to have no difficulty in seeking it out from any source.

Discussion

That Rasch calibrations provide more useful information than traditional composite scales is illustrated by a comparison of the information available using these two methods. Using composites, it's possible to distinguish people who used a feedback-seeking mode more or less frequently but not to investigate patterns of feedback-seeking. Because composite scores are ordinal, calculations of means and standard deviations and plotting of frequencies by score are not appropriate. An example of a display of the results using composites is presented in Figure 3. This figure shows boxplots of the distribution of scores along with the median scores for the three feedback-seeking scales used by Morrison (1993): monitoring, inquiring, and consulting. Using this information, it is possible to: 1) compare average values across scales, 2) see how the scores are distributed, and 3) see where the scores for a particular individual fall as compared to this distribution. Using the expanded item map presented

in Figure 2, not only can one determine the extent to which individual faculty sought feedback but one can also determine which feedback-seeking methods they used and how frequently.

Using composites, one loses the information on individual items. Looking at the boxplots gives the impression that there is more variation in frequency of monitoring than inquiry. Using the expanded item map, however, one can see that the differences by sources are in inquiry and not monitoring.

Using composites, it is not possible to detect items with idiosyncratic responses. One item, *consult written documents*, constituted a separate factor that was moderately related to the other two composites ($r=.37$; $p=.01$ with inquiry and $r=.22$; $p=.01$ with monitoring). However, there were no clues as to its misfit. Using Rasch, it was possible to identify the misfit of this item with the other two feedback-seeking modes.

Finally, using composites to measure feedback-seeking, it is not possible to see how monitoring and inquiry are related to each other. The factor analysis performed by Morrison (1993) identified three factors but the correlations among these two composite scores indicate that these are not independent factors ($r=.36$; $p=.01$). The resulting composite scores also had weak reliability (.60 and .65). The Rasch analysis clearly shows one feedback-seeking construct with good reliability (.80). Using Rasch, it is obvious that monitoring and inquiring are not polar opposites as Morrison's results suggest. If the relationship between monitoring and inquiry had been examined in previous results, it would have been obvious that the polars were no feedback-seeking and frequent feedback-seeking. The plot of monitoring and inquiry composites is presented in Figure 4. This plot confirms the findings of the Rasch analysis concerning the polars. A regression analysis performed on these data shows an intercept of approximately 1.5 and a slope of approximately 1. These results indicate that faculty monitor more frequently than inquire and this relationship is maintained as frequency of feedback seeking increases. Thus Rasch analysis provided the same information without having to conduct a statistical analysis.

Conclusions

In addition to confirming the conclusions from previous research that newcomers are more likely to monitor behavior than make inquiries of others, the results of this study show that:

1) faculty can be distinguished by their feedback-seeking mode: few do not seek feedback at all; some seek feedback only through monitoring of the behavior of colleagues; the majority of faculty seek feedback from a variety of sources, mainly from monitoring but also to some extent from inquiries of other faculty or administrators; and few faculty frequently seek feedback from multiple sources

2) the more faculty seek feedback, the more frequently both monitoring and inquiry occurs

3) information on the relationship between monitoring and inquiry and on the frequency of feedback seeking from individual sources is more than could be obtained by using traditional composite scores

Figure 1. Original Item Map of Faculty Feedback-Seeking Behavior

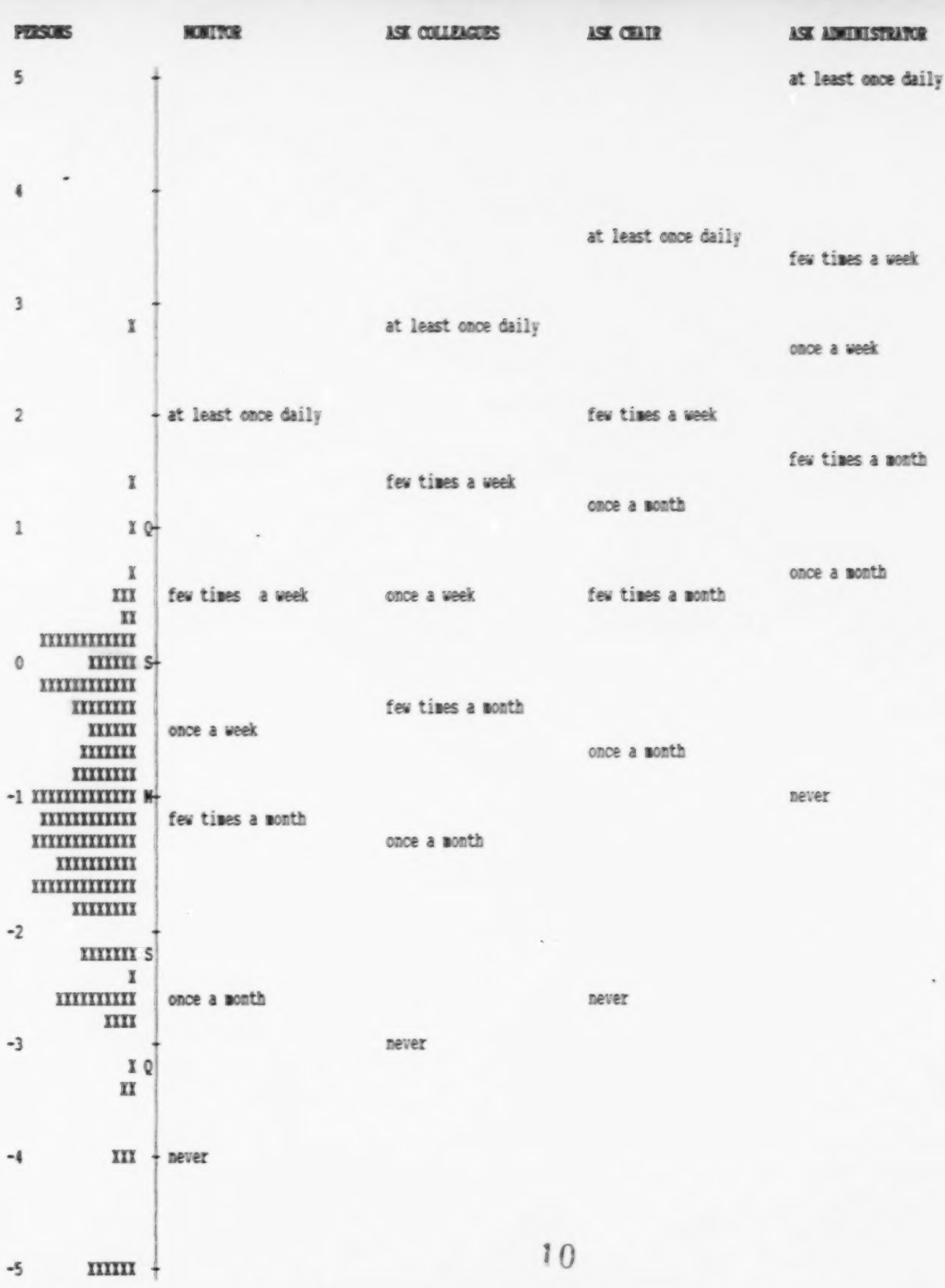

Figure 2. Expanded Item Map of Faculty Feedback-Seeking Behavior

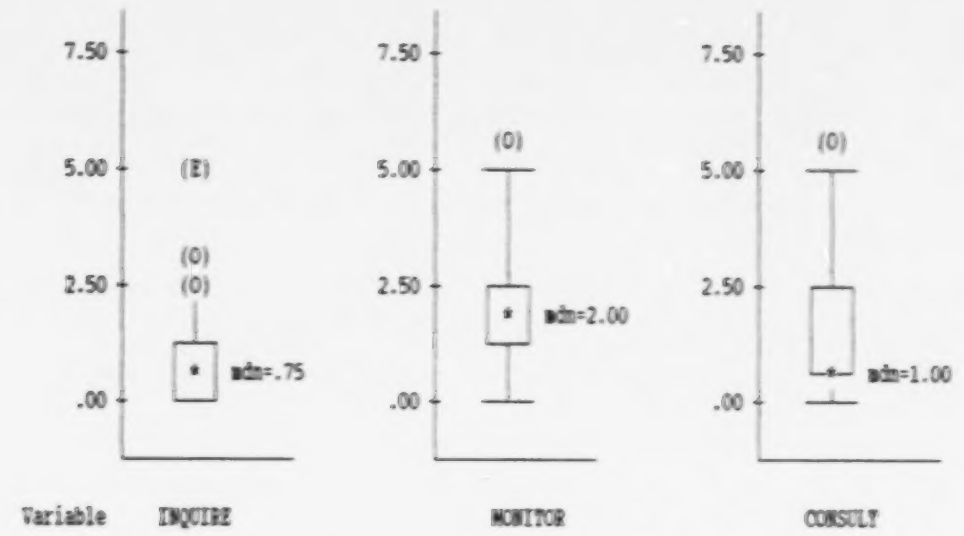

Figure 3. Distribution of Feedback-Seeking Composite Scores

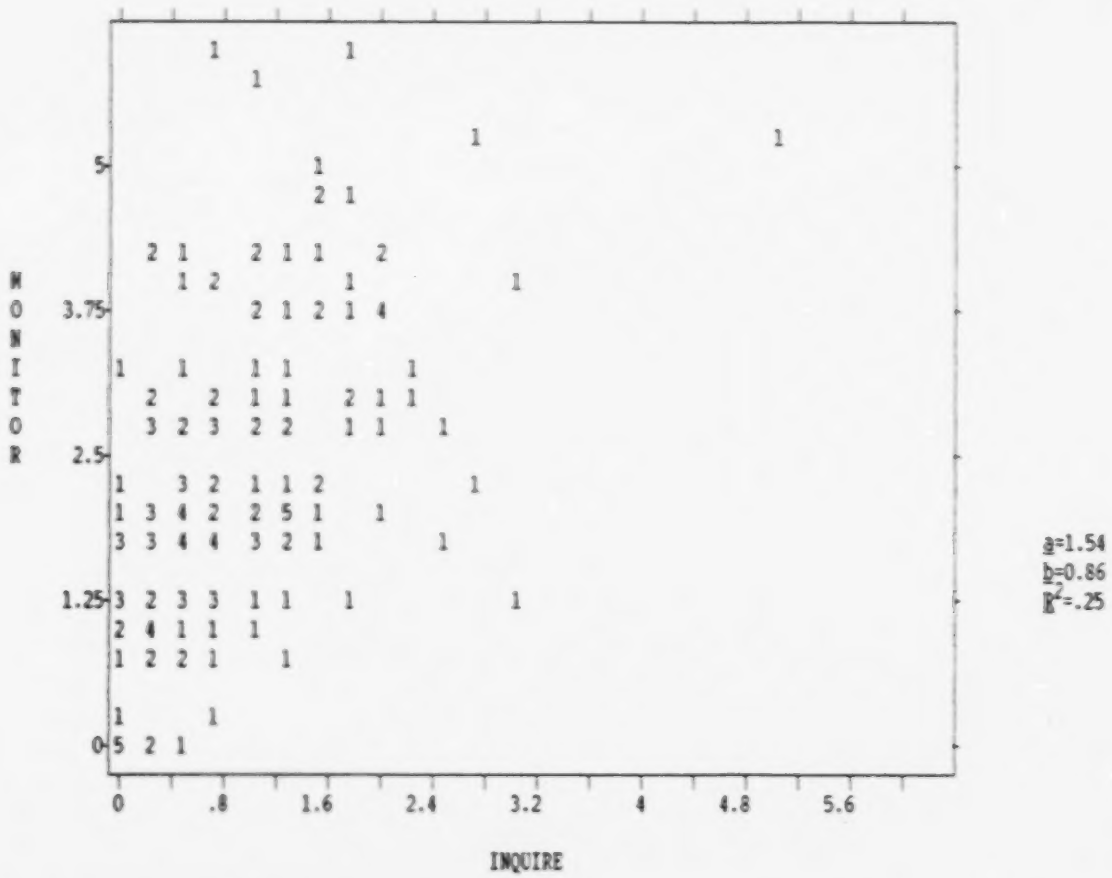

Figure 4. Plot of Monitor and Inquire Composite Scores

11

BEST COPY AVAILABLE

References

Linacre, J.M. & Wright, B.D. (1992). A User's Guide to BIGSTEPS: Rasch-Model Computer Program. Chicago: MESA Press.

Morrison, E.W., (1993). Newcomer information seeking: Exploring types, modes, sources, and outcomes. *Academy of Management Journal*, 36(3), 557-589.

Menges, R.J., Bode, R.K., Reyes, E., & Letwat, D. (1996). Seeking and using feedback about teaching: A study of junior faculty. Evanston IL: Center for the Teaching Professions, Northwestern University.

Wright, B.D., & Masters, G.N. (1982). *Rating Scale Analysis: Rasch Measurement*. Chicago IL: MESA Press.